

Incidence Angle Normalization of Radar Backscatter Data

Iliana E. Mladenova, Thomas J. Jackson, *Fellow, IEEE*, Rajat Bindlish, *Senior Member, IEEE*, and Scott Hensley, *Senior Member, IEEE*

Abstract—The National Aeronautics and Space Administration's (NASA) proposed Soil Moisture Active Passive (SMAP) satellite mission (~ 2014) will include a radar system that will provide L-band multi-polarization backscatter at a constant incidence angle of 40° . During the pre-launch phase of the project, there is a need for observations that will support the radar-based soil moisture algorithm development and validation. A valuable resource for providing these observations is the NASA Jet Propulsion Laboratory Uninhabited Aerial Vehicle Synthetic Aperture Radar (UAVSAR). However, SMAP will observe at a constant incidence angle of 40° , and UAVSAR collects data over a wide range of incidence angles (25° – 60°). In this investigation, a technique was developed and tested for normalizing UAVSAR data to a constant incidence angle. The approach is based on a histogram matching procedure. The data used to develop and demonstrate this approach were collected as part of the Canadian Soil Moisture Experiment 2010 (CanEx-SM10). Land cover in the region included agriculture and forest. Evaluation was made possible by the acquisition of numerous overlapping UAVSAR flight lines that provided multiple incidence angle observations of the same locations. Actual observations at a 40° incidence angle were compared to the normalized data to assess performance of the normalization technique. An optimum technique should be able to reduce the systematic error (Bias) to 0 dB and to lower the total root mean square error (RMSE) computed after correction to the level of the initial residual error (RMSE_{res}) present in the data set. The normalization approach developed here achieved both of these. Bias caused by the incidence angle variability was minimized to ~ 0 dB, whereas the residual error caused by instrument related random errors and amplitude fluctuations due to ground variability was reduced to approximately 3 dB for agricultural areas and 2.6 dB for forests; these values were consistent with the initial RMSE_{res} estimated using the un-corrected data. The residual error can be reduced further by aggregating the radar observations to a coarser grid spacing. The technique adequately adjusted the backscatter over the full swath width irrespective of the original incidence angle, polarization, and ground conditions (vegetation cover and soil moisture). In addition to providing a basis for fully exploiting UAVSAR (or similar aircraft systems) for SMAP algorithm development and validation, the technique could also be adapted to satellite radar systems. This normalization approach will also be beneficial in terms of reducing the number of

flight lines required to cover a study area, which would eventually result in more cost-effective soil moisture field campaigns.

Index Terms—Backscatter, incidence angle effect, incidence angle normalization, Soil Moisture Active Passive (SMAP).

I. INTRODUCTION

AS PART of its Decadal Mission program, the National Aeronautics and Space Administration (NASA) is developing a new satellite called Soil Moisture Active Passive (SMAP)¹[1] that will utilize a combined L-band radiometer (1.4 GHz) and radar (1.26 GHz) system to provide global soil moisture and freeze-thaw products. These instruments will utilize a conically scanning mesh reflector antenna with a constant incidence angle of 40° . Three different soil moisture products will be developed; passive-only with a 36-km grid, radar-only with a 3-km grid, and a combined (active-passive) product with a 9-km grid spacing. The passive-only product has heritage from the Advanced Microwave Scanning Radiometer instruments and the Soil Moisture and Ocean Salinity (SMOS) mission. However, radar-based global soil moisture retrieval at a 3-km scale and using a combined algorithm are new approaches and will require observational database to develop and validate algorithms. In addition to soil moisture, a 3-km radar-based freeze/thaw state product will offer the opportunity to discriminate between frozen and non-frozen soils.

During the pre-launch phase of SMAP, obtaining data sets for a wide range of conditions that simulate SMAP radar and combined radiometer/radar products, in conjunction with ground-based measurements of soil and vegetation conditions, is a high priority. Aircraft-based systems are a valuable resource for collecting the needed data. The only combined active-passive aircraft-based simulator that has been employed in the past is the Passive and Active L-band System (PALS), which provides only a single footprint along a flight line [2]–[4].

For the radar-only soil moisture algorithms, another aircraft option is the Uninhabited Aerial Vehicle Synthetic Aperture Radar (UAVSAR)² [5], [6]. UAVSAR observes the ground at a wide range of incidence angles ($\sim 20^\circ$ – 65°) whereas SMAP will operate at a fixed surface incidence angle (θ_i) of 40° . Incidence angle has a significant effect on backscatter, therefore, in order to fully exploit an instrument like UAVSAR, we must correct or normalize the full UAVSAR swath to the single incidence angle that SMAP will employ.

Manuscript received January 27, 2012; revised April 26, 2012; accepted May 15, 2012. Date of publication August 9, 2012; date of current version February 21, 2013.

I. E. Mladenova and T. J. Jackson are with the Hydrology and Remote Sensing Laboratory, U. S. Department of Agriculture, Beltsville, MD 20740 USA (e-mail: Iliana.Mladenova@ars.usda.gov; thomas.jackson@ars.usda.gov).

R. Bindlish is with the Science System and Applications Inc., Lanham, MD 20706 USA (e-mail: rajat.bindlish@ars.usda.gov).

S. Hensley is with the National Aeronautics and Space Administration, Jet Propulsion Laboratory, Pasadena, CA 91109 USA (e-mail: scott.hensley@jpl.nasa.gov).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TGRS.2012.2205264

¹<http://smap.jpl.nasa.gov/>

²<http://uavsar.jpl.nasa.gov/>

One approach to dealing with incidence angle effects is to use only a narrow strip from the full UAVSAR swath. Examining the backscatter sensitivity within small incidence angle increments is a common way of analyzing multi-incidence angle SAR data [7]–[9]. Typically, it is assumed that when the incidence angle range is small, the backscatter variability due to θ_i is negligible. Attempts to utilize a wider range of incidence angles have included approaches that use a cosine adjustment and other empirically based (site and sensor specific) methods [10]–[13].

Here, we propose a normalization approach that is based on a histogram matching procedure, which will be described in detail. Results will be evaluated using data collected as part of the Canadian Soil Moisture Experiment (CanEx-SM10), which was conducted during the month of June 2010 in Saskatchewan, Canada [14]. UAVSAR data were collected on multiple dates over agricultural and forest land cover conditions. The experiment design included seven overlapping swaths that were intended to cover the entire study domain at $40^\circ \pm 5^\circ$ (incidence angle variability effects would be assumed negligible over this $\sim 10^\circ$ range). This flight line design also provided coverage of the same ground location at multiple incidence angles. This aspect facilitated the analysis of the normalization technique.

Our main goal is to develop a normalization technique that will enable us to adjust the multi-incidence UAVSAR data to a fixed angle of incidence. As pointed out, these efforts are carried out in the context of the SMAP mission. The approach needs to be robust, accurate, easily applicable, and transferable. Therefore, we will assess the algorithm performance in terms of sensitivity to ground conditions, polarization, initial incidence angle, and choice of reference angle. The ultimate task is to generate a continuous map of the area by merging the individual flight lines. The comprehensiveness of the radar data set offers us the opportunity to not only evaluate the adjusted values against actual observed data, but also to explore the possibility to replicate the same results with fewer flight lines, both of which are unique to this study. Consequently, the proposed normalization may lead us to more cost-efficient experiment designs by reducing the number of lines to cover a large domain. It will also allow the full utilization of previous and future UAVSAR, and similar aircraft instruments, data sets for SMAP algorithm development and validation.

A detailed description of other incidence angle correction techniques and the approach proposed here is provided in Section II. Section III presents an overview of the study area and data sets. Results and error analysis are presented and discussed in Section IV, followed by a summary of the main findings and conclusions in Section V.

II. METHODOLOGY

In general, the amount of the transmitted microwave energy that is returned to the radar receiver is determined by the 1) target properties and 2) the system configuration. The first set of factors 1) includes topography (i.e., slope, aspect), surface roughness, canopy cover, and soil permittivity. The combination and interaction of these factors within the radar footprint can lead to additional vertical and/or horizontal heterogeneity,

which can further impact the intensity/amplitude of the return signal. The radar angular configuration (i.e., look and incidence angles), polarization, and frequency/wavelength are the major instrument-related parameters 2). Modeling the incidence angle contribution to the total backscatter variability, and isolating the effect of the other contributing factors is not a simple task. For example, assuming a specific polarization and frequency, and flat terrain, an increase in backscatter over a uniformly vegetated area could be caused by a decrease in incidence angle and/or increase in soil moisture; isolating the effects of the latter will require removing the angular dependence of the radar data. Furthermore, a varying incidence angle encumbers the direct intercomparison of data from different systems that have different viewing geometries. The commonly used approaches to normalization and the new methodology we are proposing are described in the following sections.

A. Overview of Available Normalization Techniques

One of the earliest and most often used techniques is cosine correction, which was developed by [15] and described in [10]. This method is derived from the Lambert's law for optics. The model is based on the assumption that the amount of power that is re-radiated in the upper hemisphere follows a cosine law [10]; and furthermore, because the radiation variability as a function of the observed area is also cosine dependent, the measured backscatter, $\sigma_{\theta_i}^\circ$, is related to the cosine square of the incidence angle as follows:

$$\sigma_{\theta_i}^\circ = \sigma_0^\circ \cos^n(\theta_i) \quad (1)$$

where $n = 2$ and σ_0° is the backscatter independent of incidence angle.

Consequently, the radar response, $\sigma_{\theta_i}^\circ$, acquired at an incidence angle, θ_i , can be normalized to approximate the system response at any other angle, θ_{ref} , using

$$\sigma_{ref}^\circ = \frac{\sigma_{\theta_i}^\circ \cos^n(\theta_{ref})}{\cos^n(\theta_i)} \quad (2)$$

where ref stands for reference angle, which would be 40° for SMAP.

Reference [16] noted that n is roughness dependent and suggested that the cosine approach can be improved by adjusting n using a linear regression model. The power index, n , is defined as the slope of a linear fit between $\ln(\sigma_{\theta_i}^\circ)$ and $\ln(\cos \theta_i)$. This approach was applied by the authors for each vegetation class encountered in their study area using Phased Array type L-band Synthetic Aperture Radar (PALSAR) HH-pol data acquired over the Orinoco region, Columbia. The resulting adjusted- n values varied between 0.2 and 3.4 depending on vegetation type and season, being the highest for savanna-dominated areas. It should be noted that among the five vegetation types considered by the authors, four forested classes and savannas, the savannas backscatter values were the most affected by the incidence angle. The authors concluded that the cosine-based normalization (COS) is suitable if applied over areas with "simple structure."

However, the routinely applied values of n (1 and 2) may not be sufficient to adequately account for the angle induced variability [16].

Another approach to normalization is associated with applying radiative transfer-based models. One such example can be found in [16]. In addition to the cosine normalization discussed previously, the authors approximated the ALOS PALSAR response at nadir using an extended water-cloud model. This model was further refined with the Newton optimization algorithm that was used to estimate the required model parameters. Modeled values showed good agreement with the observed data ($R^2 > 0.77$), and, as concluded by the authors, the water-cloud technique offered a better way to understand the backscatter behavior over complex ground conditions, such as dense flooded forests, than the cosine approach. However, backscattering models are complex, applicable to specific ground conditions, and require a large number of ancillary parameters [17]–[19]. Because it is impossible to obtain the required parameters either routinely or over large domains, these models are often over simplified to rely on fixed values or rough estimates and utilize extensive iterative optimization techniques. This in turn restricts their applicability to specific vegetation classes, soil moisture, and roughness conditions. For example, the empirical backscattering model of Oh *et al.* is best suited for bare soil, soil moisture (SM) ranging between 0.09 and 0.31 m^3/m^3 and has the following roughness limitations: $0.1 < ks < 6$ and $2.5 < kl < 20$ (s —RMS height; l —correlation length; k —wavenumber) [17]; the Dubois model gives good results for $SM \leq 0.35 m^3/m^3$, $\theta_i \geq 30^\circ$, low vegetation amounts ($NDVI < 0.4$), and $ks \leq 2.5$ [18]; etc.

A third type of normalization methods is based on statistical techniques. Our review identified two approaches; 1) methods that make use of linear or second order regression equations [11], [12], [20]–[22] and 2) techniques that are based on histogram/frequency matching [23], [24]. The major limitation of (1) is that the regression equations are not transferable between sensor systems and sites. Regarding (2), the histogram-based equalization approaches have been extensively used in the area of data assimilation for removing systematic differences between a model and an independent data set in terms of mean and variance; however, their application to angular correction of microwave data has been limited. The main advantage of the frequency/histogram-based techniques is that they require no or little prior knowledge, are not sensor or site specific, and most importantly, they can account for the nonlinear nature of the $\sigma^\circ - \theta_i$ relationship. The linear assumption, which is often adopted in the regression-based techniques, may be a valid approximation only if the incidence angle range is relatively small.

B. Proposed Approach

References [23], [24] scaled AirSAR (the predecessor to UAVSAR) data acquired over a savanna region located in Northern Australia using the following frequency model:

$$cdf_{ref}(\sigma_{norm_{x,y}}^\circ) = cdf_{\theta_i}(\sigma_{act_{x,y}}^\circ) \quad (3)$$

where ref stands for reference angle; $\sigma_{norm_{x,y}}^\circ$ and $\sigma_{act_{x,y}}^\circ$ represent the normalized backscatter and the backscatter observed at θ_i of each pixel, respectively, (x, y) . In this equation, the cumulative distribution functions (cdf) were generated by numerically sorting the backscatter values for each 1° incidence angle separately. The normalized image was then created by replacing the $\sigma_{act_{x,y}}^\circ$ with the $\sigma_{ref_{x,y}}^\circ$ that has the same rank as $\sigma_{act_{x,y}}^\circ$. It was assumed that each 1° line has a similar range and distribution of land cover types.

In the present investigation, we are proposing a correction approach that is based on histogram matching as opposed to the normalization with cumulative frequencies. The main difference between the two techniques is that in the frequency-based approach, the process of “scaling” the normalized lines will be “forced” to have the same dynamic range as the range of the reference line. Thus, the histogram-based (HIST) technique will be able to preserve the natural signal variability as opposed to the CDF where it is going to be limited by and to the variability observed in the reference line. Both of these approaches assume that each 1° incidence angle line adequately captures the full range of conditions observed on the ground (i.e., roughness, soil moisture, etc.). This assumption was validated using the ground soil moisture and roughness data collected during the campaign. An F -test indicated that the soil moisture and roughness variances computed per vegetation class were not different than the overall variances observed over the study areas.

In the HIST method, the multi-incidence angle UAVSAR radar data were modified to approximate the response at 40° by utilizing the lowest two central moments, the mean and variance as follows:

$$\sigma_{norm_{P,VC}}^\circ = \bar{\sigma}_{ref_{P,VC}}^\circ + \hat{\sigma}_{ref_{P,VC}}^\circ \frac{(\sigma_{act_{P,VC}}^\circ - \bar{\sigma}_{act_{P,VC}}^\circ)}{\hat{\sigma}_{act_{P,VC}}^\circ} \quad (4)$$

where σ° is radar backscatter in [dB]; “—” and “^” indicate mean and standard deviation computed per vegetation class (VC) and separately for each 1° incidence angle line; $norm$, ref and act stand for normalized, referenced and actual, and P is polarization. Data lower than the noise equivalent values (NEV) suggested by the UAVSAR team³ and greater than 3 dB were excluded when computing the mean and the standard deviation. Once these statistics were obtained, (4) was applied on per pixel basis (x and y were omitted from the formula for simplicity). Noise was most noticeable along the near- and far-range portions of the swath. However, the overall number of pixels that did not pass the specified thresholds was minimal and was associated mostly with areas observed at incidence angles greater than $\sim 60^\circ$. The flagged data were stored in an additional image, which was later used as a mask.

III. DATA SETS

Evaluation of the normalization technique proposed here was done using UAVSAR data collected during CanEx-SM10. The

³<http://uavsar.jpl.nasa.gov/instrument.html>

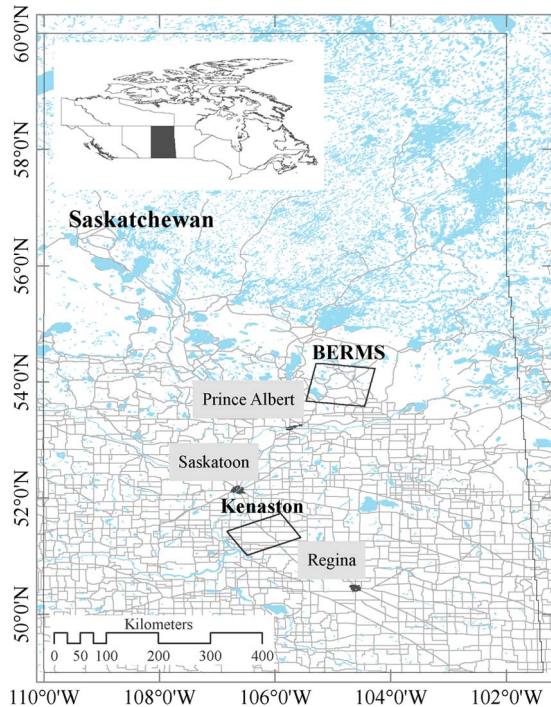


Fig. 1. Schematic representation of the CanEx-SM10 study areas, BERMS and Kenaston, located NW of Prince Albert and SE of Saskatoon, respectively. Note that the boxes show the area covered by the UAVSAR instrument during the campaign. Thus, the study area coordinates provided in the CanEx-SM10 experiment plan, which outline the two focus areas, are different and cover a slightly smaller area.

following sections describe the site, experiment design, and UAVSAR parameters.

A. CanEx-SM10 Study Area

The CanEx-SM10⁴ was conducted during the month of June, 2010, in Saskatchewan, Canada [14]. The main objectives of the field campaign were to support calibration and validation of the recently launched SMOS and algorithm development and validation for SMAP. Ground sampling (including soil moisture, temperature, vegetation, and surface parameters) and aircraft measurements were collected over two study sites, Kenaston, southeast of Saskatoon, and the Boreal Ecosystem Research and Monitoring Sites (BERMS) located northeast of Prince Albert (Fig. 1). The area is characterized by a relatively dry continental climate, and large variations between the seasonal average temperatures (climate daily max. $T_{Jul} = 24.0$ °C and min. $T_{Jan} = -23.8$ °C) and precipitation variability with the majority of the precipitation occurring during the summer season (average annual rainfall 350 mm for Kenaston and 424 mm for BERMS).⁵ The two domains exhibit relatively low terrain variability in terms of slope (Kenaston $\sim 1.55^\circ$ and BERMS $\sim 2.31^\circ$) and elevation (Kenaston $\sim 612(\pm 27)$ m and BERMS

$\sim 32(\pm 57)$ m).⁶ Approximately 60% of the two domains have elevations within \pm one standard deviation (second number) of the mean; slope is below the domains average value over $\sim 2/3$ of the area. As discussed in the beginning of Section II, topographic attributes such as slope or aspect are of particular importance for SAR observations since they can contribute to changes in the backscatter signal and need to be taken into account when calculating the incidence angle for sloping terrain, referred to as local incidence angle [25], [26]. However, given that the domains lacked significant slope-induced variability, i.e., overall, the terrain can be characterized as flat, it was assumed that incidence angle and local incidence angle are the same.

Even though the two domains have similar relief, Kenaston and BERMS differ significantly in terms of vegetation types and land use (Fig. 2). BERMS is a part of the Canadian boreal plain ecozone, and $\sim 65\%$ of the study box is occupied by forest vegetation, including jack pine, black and white spruce, aspen, birch, and tamarack. Kenaston has chernozemic loamy soils that are typical of the southern Prairie ecozone and make this area suitable for agricultural crops, pastures, and grasslands.

Vegetation information over the two domains was extracted from the GeoBase, circa-2000 land cover database.⁷ The classification was done using Landsat-5 and Landsat-7 ortho-images acquired over the 1996–2005 time period. However, 80% of the images used for the classification were collected between 1999 and 2001.

B. UAVSAR Characteristics and Data Set Description

UAVSAR is a fully polarimetric, L-band (1.26 GHz) synthetic aperture radar that evolved from its predecessor AirSAR [5], [6]. The instrument was mounted on an external pod on a NASA Gulfstream-III aircraft and flew at a nominal altitude of 12.5 km. The instrument looks to the left of the flight heading. It has a swath width of approximately 22 km and look angle range of $\sim 45^\circ$ varying between 20° in the near and $\sim 65^\circ$ in the far range. We used the calibrated UAVSAR backscatter data, which were post processed to a regular grid with slant post spacing of 7.2 m in azimuth and 5 m in range and stored in linear power units. Note that when presenting and discussing our results, the term backscatter coefficient refers to the log transformed sigma-naught value.

CanEx-SM10 was designed to capture varying soil moisture and to collect aircraft observations that simulated SMAP temporal coverage over Kenaston. As a result, data were collected with the UAVSAR on seven days (June 2nd, 5th, 6th, 9th, 13th, 14th, 15th). Only one day was obtained over BERMS (June 16th).

As mentioned earlier, in an attempt to simulate SMAP constant incidence angle data using UAVSAR and considering the tradeoff between the costs of flight lines and variability introduced by angular range, the seven flight lines were collected with the same heading and significant overlap. The alignment

⁴<http://pages.usherbrooke.ca/canexsm10/home.php>

⁵National Climate Data and Information Archive, Environment of Canada (www.climate.weatheroffice.gc.ca); the climate averages were computed using stations with data record of at least 15 years between 1971 and 2000. Temperature and precipitation data presented in the text are based on two stations located at $53^\circ 13'N$, $105^\circ 40'W$ (Prince Albert, BERMS site) and $52^\circ 10'$, $106^\circ 43'W$ (Saskatoon, Kenaston site).

⁶Slope and elevation values were computed using the Shuttle Radar Topography Mission-derived Digital Elevation Model.

⁷<http://www.geobase.ca/geobase/en/data/landcover/index.html>

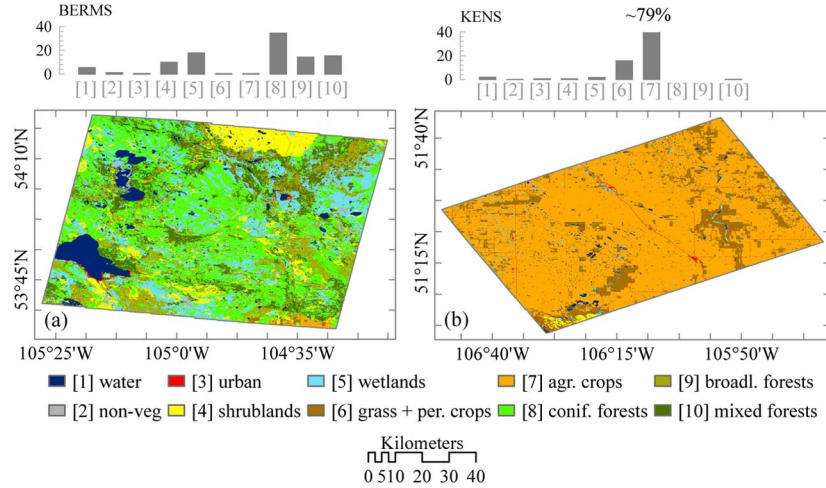


Fig. 2. Land cover map of BERMS (a) and Kenaston (b) based on circa-2000 land cover database.

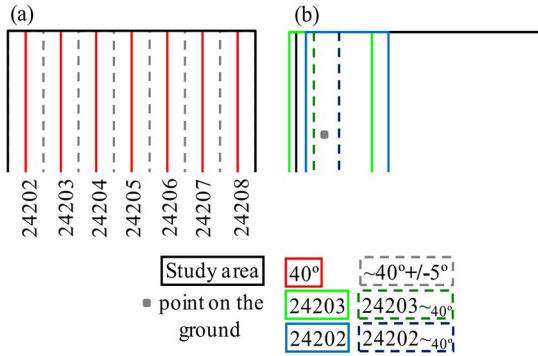


Fig. 3. Schematic representation of the UAVSAR flight lines. The gray square represents an observation point on the ground. Plot (a) shows the seven subsets centered on 40° that correspond to each of the individual flight lines and plot (b) illustrates an area on the ground observed by two flight lines at different incidence angles.

and spacing were chosen to provide full coverage of each of the two domains at an incidence angle of $40^\circ \pm 5^\circ$ (see Fig. 3, plot a). Each 10° portion corresponds to an approximately 4-km section shifted ~ 4.4 km from the near range (assuming average aircraft altitude of 12.5 km).

IV. ANALYSIS

The wide range of incidence angles of the UAVSAR, combined with the vegetation diversity of the CanEx-SM10 study areas (see Section III-A.1) and the varying moisture conditions during the campaign, provided a data set suitable for development of normalization techniques and potentially retrieval/downscaling algorithms. Furthermore, because of the extensive overlap of the flight lines, the same point on the ground was observed at different incidence angles. Thus, the data set offers the opportunity to evaluate the normalized output with actual data observed at 40° .

In addition to the proposed HIST approach, the UAVSAR data were normalized using two of the other techniques discussed earlier: the COS and the *cdf* equalization approach (CDF). The CDF approach was applied as described in [23] and [24]. However, here, the *cdfs* were built using the

same UAVSAR σ° range as for the histogram normalization (i.e., $\text{NEV} < \sigma^\circ < 3$ dB). The assumption of uniform vegetation variability was validated by comparing the frequency distributions of the land cover types encountered in each 1° incidence angle line. Percent difference per land cover type was on average $\leq 2\%$ for the major and $< 0.5\%$ for the minor land cover types. This confirms that for the CanEx-SM10 domains, it is safe to adopt the above assumption.

As pointed out in Section III-A-2, the flight lines were specifically designed so that the domains would be entirely covered at the nominal SMAP incidence angle ($40^\circ \pm 5^\circ$). The large overlap between the flight lines provided multiple observations of the same point on the ground at different incidence angles (Fig. 3, plot b). Typically, the accuracy of incidence corrected SAR data are assessed qualitatively by visual comparisons between the normalized and original image or by evaluating the accuracy of products derived using the normalized data such as land cover or soil moisture [12], [13], [20], [23]. The unique design of the UAVSAR data collection in CanEx-SM10 offered the opportunity to assess performance using data actually observed at 40° .

The seven flight lines were flown in the following order: 24202 first and 24208 last (Fig. 3, plot a). Thus, there are seven transects centered on $40^\circ (\pm 0.02^\circ)$ within the study area. When referring to an area observed at an angle θ_i in flight line “xxxxx” the following notation was adopted: $\theta_{i\text{xxxx}}$, where $\theta_i (\pm 0.02^\circ)$. Depending on the location of the flight line throughout the domain, each 40° transect overlaps with between 1 and 4 other flight lines; for example, the area observed at 40°_{24202} is also observed at $\sim 23^\circ_{24203}$ (Fig. 3, plot b); the area observed at 40°_{24204} is also observed at $\sim 23^\circ_{24205}$ $\sim 51^\circ_{24203}$ and $\sim 59^\circ_{24202}$, and so on. Due to noise issues with the backscatter data, areas observed at angles greater than 60° were excluded.

Root mean square error (RMSE) and bias were used as measures of error and indicators of improvement relative to the original data measured at 40° . These statistics were computed per flight day using all available matching points (i.e., no vegetation discrimination) and also for each land cover type separately. The following section describes the notation that

will be used in the results/discussion section. As previously mentioned, the same area on the ground is observed at 40° and some other angle, θ_i . The term $\sigma_{ORG,40^\circ}^\circ$ indicates the original backscatter data actually observed for the pixel at 40° . $\sigma_{ORG,\theta_i}^\circ$ is the original backscatter data observed over the same pixel at θ_i . $\sigma_{NORM,\theta_i \rightarrow 40^\circ}^\circ$ stands for the backscatter data normalized from θ_i to 40° using one of the three approaches, HISTogram, COSine, or CDF (i.e., $NORM = HIST$, $NORM = COS$, $NORM = CDF$); consequently, the original RMSE ($RMSE_{ORG}$) and bias ($Bias_{ORG}$) were computed using $\sigma_{ORG,40^\circ}^\circ$ and $\sigma_{ORG,\theta_i}^\circ$, while $RMSE_{HIST}$, $RMSE_{CDF}$, and $RMSE_{COS}$ were computed using $\sigma_{ORG,40^\circ}^\circ$ and $\sigma_{NORM,\theta_i \rightarrow 40^\circ}^\circ$.

Using the above definitions, RMSE and Bias were computed using the following formulas:

$$RMSE = \sqrt{\left[\left(\sigma_{ORG,40^\circ}^\circ - \sigma_{ORG,\theta_i/NORM,\theta_i \rightarrow 40^\circ}^\circ \right)^2 \right]} \quad (5)$$

$$Bias = \left(\sigma_{ORG,40^\circ}^\circ - \sigma_{ORG,\theta_i/NORM,\theta_i \rightarrow 40^\circ}^\circ \right) \quad (6)$$

where σ° is radar backscatter in [dB] and “—” indicates mean. The final error estimate is in units of [dB]. An alternative approach to computing the error (E) in [pwr] is provided in

$$E = 10 * \log_{10} \left(\frac{\sigma_{ORG,40^\circ}}{\sigma_{ORG,\theta_i/NORM,\theta_i \rightarrow 40^\circ}} \right) \quad (7)$$

where σ is radar backscatter in [pwr]. Using log transformations, it can be demonstrated that the E estimated with (7) is equivalent to the bias computed using (6).

As pointed out by [27], RMSE is a very widely used and easily understood error statistic. However, it is important to note that RMSE includes two components, systematic (bias) and residual error ($RMSE_{res}$), that are related through the following formula [27], [28]:

$$RMSE = \sqrt{Bias^2 + RMSE_{res}^2} \quad (8)$$

The $RMSE_{res}$ is defined by the standard deviation of the errors. It explains the random component of the errors ($RMSE_{RANDOM}$) and the component related to the bias in the amplitude of fluctuations ($Bias_{AMPLITUDE}$), where $RMSE_{res} \approx RMSE_{RANDOM} + Bias_{AMPLITUDE}$. For a radar system, the $RMSE_{RANDOM}$ is typically related to random noise, speckle, and instrument precision. The random error can be computed via the following formula:

$$RMSE_{RANDOM} = \frac{1}{\sqrt{N}} \left(1 + \frac{1}{SNR} \right) \quad (9)$$

where SNR is signal to noise ratio and N is the number of looks or pixels averaged together to reduce the speckle and thermal noise. For the UAVSAR data, $N = 36$ and the $SNR > 15$ dB, and thus backscatter variation is typically dominated by speckle noise. Overall, the $RMSE_{RANDOM}$ is normally inherent

and hard to correct. On the other hand, the amplitude bias by definition has 0 mean, but can have persistent season/spatial variations. In our case, these fluctuations can be caused by the soil moisture and/or surface roughness variability within each vegetation class. The amplitude bias is expected to be reduced, if not completely removed, by the normalization approach via the standard deviation correction. Between the two $RMSE_{res}$ components, the $RMSE_{RANDOM}$ is independent of ground influences suggesting that it should be constant throughout the swath width and between lines. Adequate removal of the amplitude bias errors caused by the aforementioned additional ground factors would result in constant $RMSE_{res}$ regardless of initial incidence angle.

The mean bias (bias) is predictable and often a result of some systematic causes, e.g., antenna pattern knowledge errors, differences in ground conditions, etc. From this point on unless specified differently, the term “bias” refers to mean bias. If these causes are adequately identified and accounted for, the bias can be easily corrected. It should be remembered that the main goal of our approach is to account for differences caused by differences in θ_i , which is systematic in nature. In an ideal system, the bias should approach 0 (i.e., observed = predicted) and $RMSE_{res}$ should approach the total RMSE [27]–[29]. Thus, a perfect normalization model should be able to reduce the bias to 0 dB and lower the $RMSE_{NORM}$ to the level of the residual error ($RMSE_{res,ORG}$). Furthermore, if the mean bias is 0 (if the correction is successful), then the final RMSE computed using the corrected data will be actually equivalent to the residual part of the total RMSE error ($RMSE_{NORM} = RMSE_{res,NORM}$).

A. Algorithm Evaluation

As a first step, we evaluated the incidence angle effect by examining the backscatter difference between the near and the far range of the swath ($\Delta \sigma_{ORG,P,VC}^\circ = \sigma_{ORG,P,VC}^{Near} - \sigma_{ORG,P,VC}^{Far}$). The $\sigma_{ORG,P,VC}^\circ$ versus θ_i response curves were derived separately for each vegetation class (VC) encountered in the two domains and per polarization (P) by calculating the average $\sigma^\circ(\overline{ORG})$ per 1° θ_i change; since the $\sigma_{ORG,P,VC}^\circ$ versus θ_i response curves were similar with the ones published in literature, we will not include any plots here [30].

The θ_i impact appeared to be smaller over the forest dominated BERMS site, where $\Delta \sigma_{ORG,P,Forests}^\circ$ difference was similar for the three polarizations (i.e., $\Delta \sigma_{ORG,HH,Forests}^\circ = 7.6$ dB; $\Delta \sigma_{ORG,VV,Forests}^\circ = 9.2$ dB; $\Delta \sigma_{ORG,HV,Forests}^\circ = 7.0$ dB; values represent an average from the three forest groups observed in the domain: coniferous, broadleaf, and mixed, where the standard deviation ranged between 0.1 and 0.5 dB depending on polarization). The similarity in the absolute value of the backscatter difference for the three forested classes, broadleaf, coniferous, and mixed, over the BERMS is not unexpected. Comparable results were reported by [31]. In that study, the authors examined the L-band radar response over a boreal forest located in Alaska, USA. They found that the mean backscatter of white spruce, black spruce, and balsam poplar (broadleaf)

measured at a 44° incidence angle were very similar. Average backscatter values computed over BERMS at the same incidence angle were consistent with the values reported by [31], i.e.,

$$\begin{aligned}\sigma_{Broadleaf_BERMS/Alaska(P)}^\circ &= -9.2/-9.5 \text{ dB (HH);} \\ &\quad -15.5/-14.5 \text{ dB (HV);} \\ &\quad -10.4/-10.8 \text{ dB (VV)} \\ \sigma_{Needle-leaf_BERMS/Alaska(P)}^\circ &= -8.9/-9.5 \text{ dB (HH);} \\ &\quad -15.8/-15.7 \text{ dB (HV);} \\ &\quad -10.7/-10.5 \text{ dB (VV)}.\end{aligned}$$

In general, observations of a forested canopy are expected to be less sensitive to incidence angle [30]. However, this was not what we observed. The large backscatter difference between the near and far range of the swath suggests spatial nonuniformity in terms of canopy density, tree structure, and height along with nonhomogeneity within the forested vegetation class. Boreal forests are typically a needle-leaf-dominated environment; nevertheless, they are characterized by sparse canopy cover and represent a mixture of broadleaf vegetation and shrubs, which is further complicated by the presence of isolated patches of grasslands, flooded vegetation, and small open water bodies. All of these factors can cause strong vertical layering and horizontal variability [32]–[35]. Under dense forested conditions, the backscatter return should be mostly a result of volume and double-bounce scattering; the ground contribution is minimal. At small incidence angles in coniferous forests with sparse canopy closure, the radar energy will reach the ground, and the reflected energy will be a function of some double-bounce and direct ground scattering from the underlying litter/canopy layer and soil. It will also be influenced by the dielectric properties of the soil and the underlying vegetation [35]. For the same canopy type and closure conditions, if the incidence angle is large, the backscatter signal will be controlled primarily by the trunk-ground and branch-ground double-bounce scattering mechanisms and the tree dielectric properties [36]. On the other hand, over dense broad-leaf forests regardless of the incidence angle, the ground contribution will be absent, and the reflected energy will be dominated by volume and double-bounce scattering. This background information explains the different sensitivities of broad- and needle-leaf forests to incidence angle variation and provides a reasonable explanation for the observed difference between the near and far range over the forest types at the BERMS site. Furthermore, similar backscatter behavior was also reported by [37], who examined RADARSAT-1 data acquired over a black spruce and jack pine dominated site located in the Northwest Territories, Canada. Additionally, it appears that the three forest classes that exist in BERMS (coniferous, deciduous, and mixed) are not homogenous enough for the UAVSAR system to pick up different responses, thus, they can be grouped together as a single class.

Approximately 1 dB greater difference compared to forests was observed over areas covered by barren land, shrublands, wetlands, and grasslands and perennial crops. The $\Delta\sigma_{ORG_{P,Agr}}^\circ$ computed over agricultural areas in BERMS were higher relative to the rest of the vegetation classes

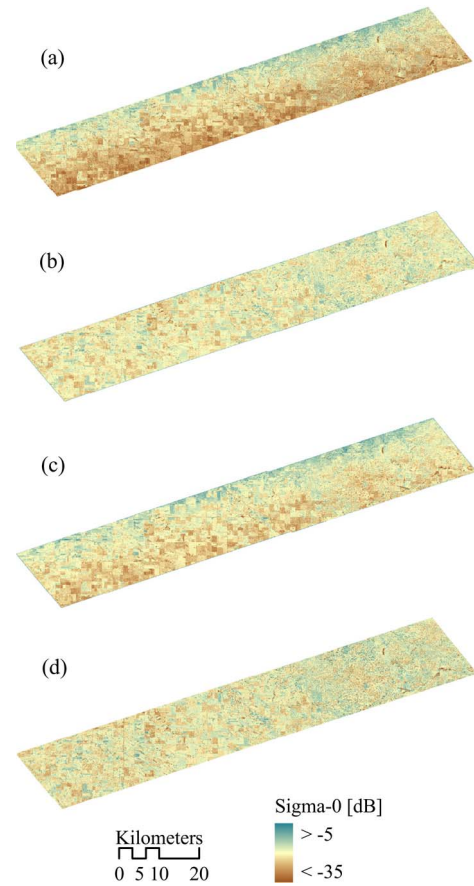


Fig. 4. Examples of UAVSAR data normalized to 40° using HISTogram (b), COSine (c), and CDF (d) normalization, while the original un-corrected backscatter is shown in plot (a). Note: Flight line: 24202, polarization: HH, date: June 2, 2010. Geographic bounds in terms of upper left and lower right corner of the flight line: 51.73° N– 105.90° W and 51.25° N– 106.69° W.

and were consistent with the values estimated over the same vegetation class in Kenaston ($\Delta\sigma_{ORG_{HH,Agr}}^\circ = 15.3/14.5$ dB; $\Delta\sigma_{ORG_{VV,Agr}}^\circ = 11.6/11.3$ dB; $\Delta\sigma_{ORG_{HV,Agr}}^\circ = 9.2/9.5$ dB; BERMS/Kenaston). Overall, the backscatter data collected over Kenaston indicated greater sensitivity to the aircraft angular variability, which is most likely due to the vegetation types encountered in this study area. The $\Delta\sigma_{ORG_{P,VC}}^\circ$ appeared to be independent of moisture change (standard deviation value computed per vegetation class using all 7 days was on the order of 0.5 dB for all three polarizations). For all vegetation types except forests, the VV $\sigma_{ORG_{VV,VC}}^\circ$ versus θ_i response curve was slightly greater than the HH-generated curve, which is consistent with literature, where the difference was small/negligible under smaller θ_i (i.e., $\theta_i < 30^\circ$) and increased as the incidence angle increased. The opposite was observed over forests, where the difference between the HH- and VV-curves was minimal, i.e., the HH-pol response was slightly greater than the VV-pol by ~ 1 dB. This reversal over forested sites was reported by other authors as well, i.e., [7] and [31]. As expected, the $\sigma_{ORG_{HV,VC}}^\circ$ was lower (10 to 15 dB) compared to the VV-polarized data.

A visual representation of the impact of incidence angle on the difference between the swath edges can be seen in Fig. 4, plot a. This particular example was generated using

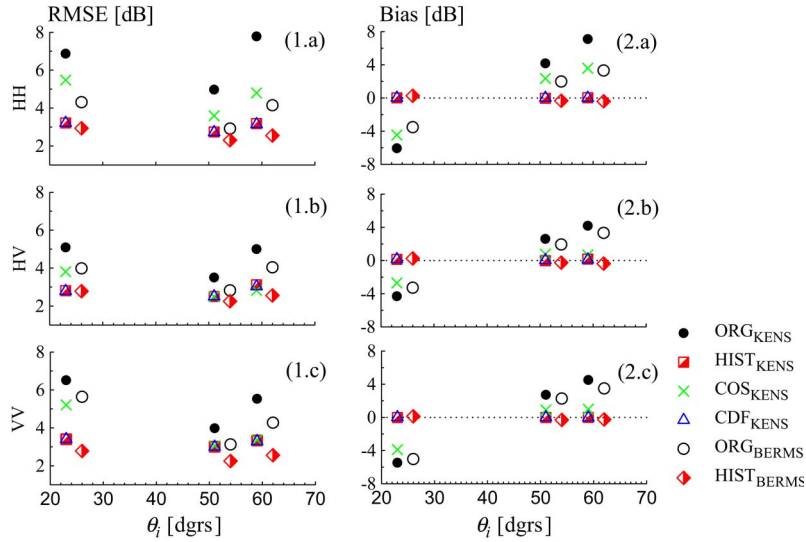


Fig. 5. Overall performance in terms of RMSE and bias of the HIST (semi-filled square), COS (cross), and CDF (triangle) approaches over Kenaston (KENS) and the HIST (semi-filled diamond) approach over BERMS. The original statistics computed against the actual 40° data are represented by the circles, where the normalized backscatter data were initially observed at $\sim 22\text{--}23^\circ$, 52° , and 59° ; for clarity, the incidence angles of BERMS were offset by 3° when plotting.

HH-polarized data obtained over Kenaston on June 2nd, 2010, when the soil moisture conditions were dry. Similar backscatter variability across the swath width, i.e., high σ° in the near and decreasing toward the far range, was observed in the other polarizations (VV and HV), and under intermittent and wet moisture conditions. As described by [23], [24] the derivative of backscatter with respect to incidence angle varies throughout the swath; under the CanEx-SM10 ground conditions, the derivative was larger up to $\sim 40^\circ$ after which it decreased.

Spatial plots of the resulting incidence corrected backscatter data using the HIST, COS, and CDF methods are shown in Fig. 4, plots b, c, and d, respectively. Among the three approaches, the COS technique had the worst performance. The incidence angle effect remained clearly visible along the near-range. A visual examination of the plots indicated that the HIST and CDF had comparable performance, which is expected given that both techniques are based on similar statistical moment matching. Error analysis indicated that there is only marginal difference between the two techniques (Fig. 5). As mentioned previously, the CDF approach preserves the statistical characteristics of the reference line. Therefore, it is anticipated that each 1° incidence angle line would have the same dynamic variability as the reference line. However, this is not expected for the HIST normalized data where the dynamic spread is not limited by the min and max range of the reference line. We examined the histograms (15th and 85th percentiles to avoid outliers) of the CDF- and HIST-corrected data sets over flight line 24204 (central for Kenaston) from June 15th, 2010. The 15th and 85th percentile values for the CDF-corrected data were consistent throughout the swath width while they varied for the HIST-corrected image confirming the above expected difference. These percentile values for the CDF-based image were -25.8 dB and -17.4 dB, respectively, regardless of original θ_i , whereas for HIST-based image, these percentile values were -23.2 dB and -14.9 dB, respectively, for 50° and -24.4 dB and -15.6 dB, respectively, for 60° incidence angles.

Given our general knowledge of the θ_i impact on the backscatter behavior, we would expect the bias to be negative if an area is observed at θ_i that is lower than the reference angle, and vice versa if θ_i is greater than 40° (i.e., $\mu_{60^\circ} < \mu_{40^\circ} < \mu_{20^\circ}$). This tendency is evident in the response of all three polarizations and over both study areas, where Bias_{ORG} is smaller over the forest-dominated BERMS site. Both, HIST and CDF, adequately removed the bias (Fig. 5, plots 2.a, 2.b, and 2.c) and reduced the error to ~ 3 dB and ~ 2.6 dB over Kenaston and BERMS, respectively. As expected, the final RMSE computed using the corrected data ($\text{RMSE}_{\text{HIST}}$) had a similar order of magnitude as the residual part of the RMSE_{ORG} ($\text{ubRMSE}_{\text{ORG}} = 2.94(\pm 0.3)$ dB and $2.3(\pm 0.1)$ dB for Kenaston and BERMS, respectively, where the RMSE value represents an average RMSE of the three polarizations and the value in the brackets shows the standard deviation).

Although, our main goal is to provide a data set that mimics the SMAP viewing geometry (40°), a robust normalization technique should be able to adequately correct to any other reference angle. Therefore, we evaluated the HIST sensitivity to the reference angle by normalizing the data to two additional angles, 30° and 50° . We limited the analysis to the HIST since the CDF has similar performance and the HIST technique gave more accurate results as compared to the COS technique. Results from these analyses are summarized in Fig. 6. For all three θ_{REF} scenarios (i.e., $\theta_{\text{REF}} = 30^\circ | 40^\circ | 50^\circ$), the $\text{RMSE}_{\text{HIST}}$ was reduced to ~ 3 dB regardless of the polarization and initial θ_i [Fig. 6, plot a, where $\text{RMSE}_{\text{HIST}/30^\circ} = 3.37(\pm 0.3)$ dB, $\text{RMSE}_{\text{HIST}/40^\circ} = 3(\pm 0.2)$ dB, $\text{RMSE}_{\text{HIST}/50^\circ} = 2.8(\pm 0.2)$ dB]. The standard deviation value given in the brackets shows the polarization-dependent variability. Values were in line with the results presented in Fig. 5.

Very noticeable in Fig. 5 is the large variability in RMSE_{ORG} . Our initial interpretation is that the lower the reference angle, the greater the RMSE_{ORG} . However, when interpreting these results, the impact of $\Delta\theta$, where $\Delta\theta$ is the

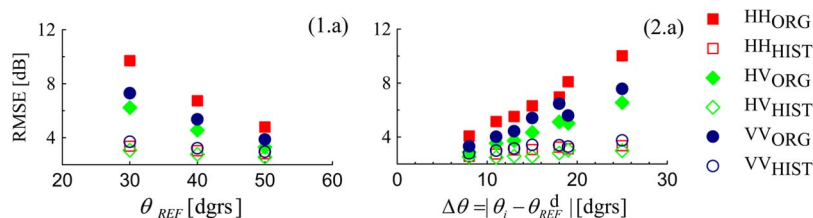


Fig. 6. Sensitivity of the HISTogram technique to reference angle. Plot (a) shows the RMSE as a function of polarization and reference angle, whereas plot (b) shows the RMSE adjustment for the three polarizations as a function of distance from the reference angle. Note: Plot summarizes results computed using only data acquired on June 2, 2010, where $\sigma_{HIST(45^\circ, 55^\circ \rightarrow 30^\circ)}^\circ$, $\sigma_{HIST(22^\circ, 51^\circ, 59^\circ \rightarrow 40^\circ)}^\circ$ and $\sigma_{HIST(37^\circ, 58^\circ \rightarrow 50^\circ)}^\circ$.

absolute difference between θ_i and θ_{REF} , needs to be taken into account; the latter is illustrated in plot b of Fig. 6 (for original θ_i see note under Fig. 6). As it can be seen the RMSE improved the least when this difference was smallest and increased with $\Delta\theta$. In addition, it should be noted that the adjustment is minimal for $\Delta\theta < 10^\circ$ (Fig. 6, plot b). To a certain extent, this would support the common assumption that incidence angle effects can be ignored if the incidence angle range is relatively small. However, taking a closer look at the original bias computed relative to the original 40° -observed data ($Bias_{ORG} = 2.3$ dB, average between the three polarizations), it clearly shows that there is incidence angle-induced variability. Furthermore, if not taken into account, this difference would result into stripping when trying to merge the individual flight lines into a continuous map.

In summary, the accuracy of the HIST normalization is not dependent on the choice of reference angle, but the RMSE improvement and the initial $RMSE_{ORG}$ are dependent on the angular distance from the reference line. Consequently, if we want to minimize the systematic errors throughout the swath width, the best reference angle will be the mid-swath incidence angle [12]. However, when comparing different SAR systems this angle is not always the same (i.e., UAVSAR versus PALSAR) and, as demonstrated, the total $RMSE_{HIST}$ was persistently reduced to approximately 3 dB, which is consistent with the initial residual error, irrespectively of ground conditions and angular characteristics (Figs. 5 and 6) indicating the robustness of the histogram approach.

B. UAVSAR Composite Maps and Effect of Reduced Data Coverage

Published normalization studies have typically focused on a single flight line or satellite swath. Even though the normalized image may appear visually smooth (i.e., no obvious difference between the swath edges), often there is residual bias on both sides of the reference angle that can vary. One of our goals was to be able to merge the individual flight lines into a seamless map; the resulting composite maps of the area are shown in Fig. 7.

As anticipated, each individual line remained visible in the composite of the original data due to the incidence angle-induced variability (Fig. 7 plot a). A COS-normalized merged map (not shown here) still had stripping similar to the original composite map. This was expected given that the COS-normalized data remained biased relative to the 40° (original) backscatter data (Fig. 5). In addition, the bias was positive for $\theta_i > \theta_{REF}$ and the opposite for $\theta_i < \theta_{REF}$. The HIST

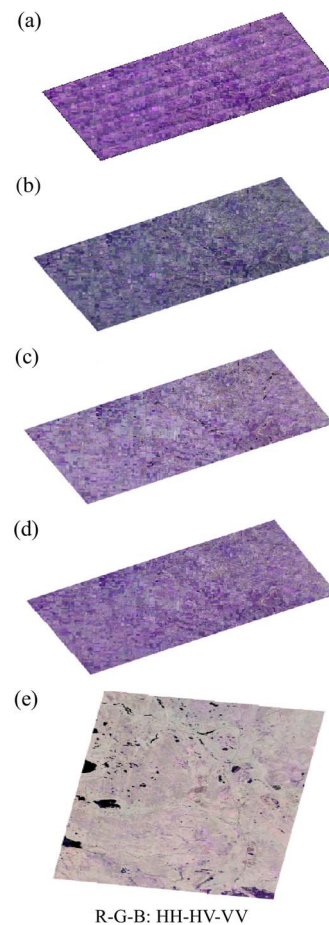


Fig. 7. Examples of mosaic maps. (a) uses the original UAVSAR data. The strips that are clearly noticeable correspond to the individual flight lines. Even though the incidence angle range of the subsets is relatively narrow, the incidence angle effect is clearly visible. Mosaic maps of the two domains generated after applying the HIST normalization are shown in plots (b) through (e). (a) KENS_ORG_June 2nd, 2011; (b) KENS_June 2nd, 2011; (c) KENS_June 9th, 2011; (d) KENS_June 13th, 2011; (e) BERMS_June 16th, 2011. Note: Data used to generate the maps are in [dB]. Refer to Figs. 1 and 2 for geographic extent.

approach was able to adequately normalize the full swath width, and the merged maps of the two domains, shown in Fig. 7 plots b–e, are continuous and do not have this incidence angle associated striping. The three Kenaston maps shown in this figure depict different moisture conditions, i.e., June 2nd—dry, June 9th—wet, and June 16th—intermittent (Fig. 7 plots b, c, and d, respectively). It appears that the performance and the accuracy of the HIST technique are independent of the soil wetness, which is also supported by the standard deviation value computed using the daily $RMSE_{NORM}$ estimates ($< \pm 0.1$ dB).

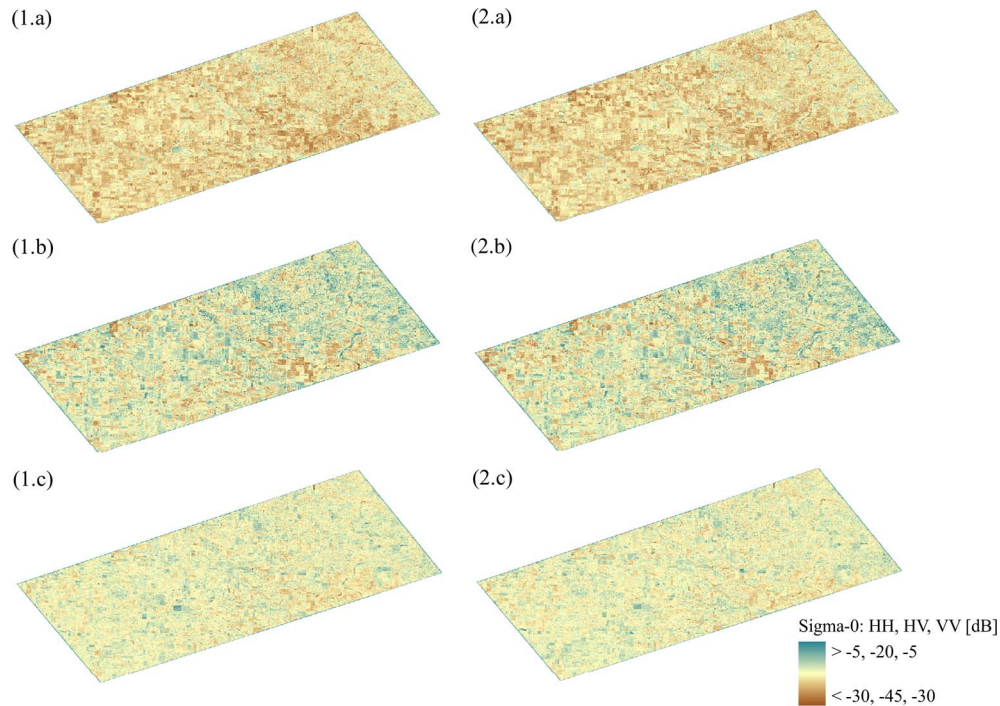


Fig. 8. Example of mosaic maps for June 2nd, 2010, using all flight lines (plots 1.a, 1.b, and 1.c, corresponding to HH, HV, and VV polarizations, respectively) and using every other flight line (plots 2.a, 2.b, and 2.c, corresponding to HH, HV, and VV polarizations, respectively). The alternative composite presented in the second column was generated using flight lines 24202, 24204, 24206, and 24208. Note: Refer to Figs. 1 and 2 for geographic extent.

Our ultimate goal is to provide the highest quality normalized data set. Therefore, we utilized all of the data available to us. The subsets used to generate the composite maps were relatively narrow, and as a result a large portion of the original individual swaths was actually disregarded. As part of our investigation, we examined if we could generate the same results by using more data from fewer flight lines. Being able to replicate the “original” composite would mean that the same area could be covered with a smaller number of flight lines, potentially leading to more cost-effective field campaigns. What we expect to find is that after normalization, the same area observed in two overlapping lines will be the “same” in terms of its adjusted backscatter value. An example is presented in Fig. 8, where the composite maps in the left column were created using all seven flight lines (“original” composite), while the maps shown in the right column were generated using four out of the seven (every other) flight lines (24202, 24204, 24206, 24208). Visual inspection shows the resulting composites are quite similar. Comparable results were obtained using only three lines (24023, 24205, and 24207) (not shown here). Results from the “alternative” composites were evaluated against the “original” composite on a daily basis as well as against the available reference data. The RMSE difference for both “alternative” scenarios was 3.0 (± 0.05) dB, 2.7 (± 0.06) dB, and 3.3 (± 0.06) dB for HH, HV, and VV-polarization, respectively (values were computed as an average of the seven daily estimates), which has a similar order of magnitude as the statistics presented and discussed earlier; note that if the same flight line was used in both composites, the overlapping area was excluded from the statistics. This means that the agreement between the two composites (i.e., “original” versus

“alternative”) over major portion of the domain will be within ± 3 dB (or less). When we examined the images as a whole (i.e., did not exclude the common flight lines), we found that 70% (60%) of the composites agree within less than ± 1 dB and 89% (85%) within ± 3 dB [percentages given here represent polarization average and 4-(3-)lines scenario]. Some of the pixels that deviated more than the RMSE difference value were classified as water, wetland, urban. In addition, as explained previously in the methodology, each flight line was normalized using its own reference histogram derived using the backscatter data observed at 40° , so part of the dissimilarity between the normalized values observed in two lines can be attributed to differences in their reference lines. However, a more significant contributor to the RMSE difference was agricultural fields with a distinct row orientation. It should be noted that the amount of vegetation cover in the fields over the duration of the campaign was minimal. Therefore, it is safe to assume that the returned energy was dominated by direct ground scattering (i.e., vegetation impact was negligible). Generally, the backscatter signal from agricultural fields varies strongly as a function of both θ_i and row orientation relative to the radar view angle [30], [38]. Over Kenaston, the agricultural fields appeared brighter, row structure was more clearly noticeable, and the backscatter return was greater when observed at lower θ_i (i.e., 25°). These characteristic differences remained even after normalization, which indicates that the reference histograms used to adjust the agricultural fields needed to be developed more carefully. As it is, the reference histograms for crops are built utilizing all the available data observed over agricultural vegetation identified in the land classification map. A possible solution would be to further sub-divide the broad agricultural

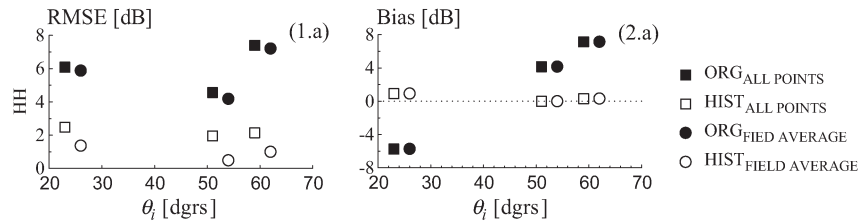


Fig. 9. RMSE and bias computed using all available points in each “homogenous” field versus statistics computed using field average values. The residual component of the RMSE caused by instrument-related errors and amplitude fluctuations due to variability in ground conditions can be significantly reduced by up-scaling the UAVSAR data; for clarity, the corresponding incidence angles of the field average-based RMSE and bias were offset by 3° when plotting.

class based on row structure (if present) and normalize each sub-class separately. Another option when generating the composites is to use a more sophisticated mosaic logic. For example, we could use the flight line that has the smallest difference relative to a sample field observed at 40° . Based on the results presented in Fig. 6 this should be the flight line in which the field in question is observed at θ_i closest to the reference angle. In addition to the aforementioned difference between the reference lines and the impact of row orientation, the RMSE difference observed between the original and the alternative composites can be attributed to the presence of standing water and flooded vegetation that were very typical over Kenaston during CanEX-SM10 as a result of frequent large rainfall events.

The analysis presented here shows that a much wider portion than the pre-defined 10° segment can be used. Given the fact that the error was successfully reduced to a value consistent with the original residual RMSE (Figs. 5 and 6, plot b) also indicates that the correction adequately adjusts the backscatter over the whole swath width. This means that reducing the number of flight lines is a viable option, which in turn may result in more efficient experiment designs.

C. Discussion of the Error Statistics

As discussed in the beginning of this section, the total RMSE includes two components, systematic and residual. Our normalization approach targets differences caused by systematic causes. Thus, it is expected that the total error after correction will be lowered to the level of the initial residual error computed using the original un-corrected data. All of the analyses presented have shown that the normalization approach is able to successfully reduce the final bias to 0 dB over both domains and to lower the total $\text{RMSE}_{\text{HIST}}$ to ~ 3 dB and ~ 2.5 dB for Kenaston and BERMS, respectively. As noted, these values are not unexpected since they are consistent with the initial residual error, which demonstrates the robustness of the proposed HISTogram approach. However, the magnitude of the final total error may be relatively large for certain applications such as soil moisture retrieval. Hence, further efforts should be focused on how we can reduce the pre-existing residual error, which would then lead to a smaller total error. Note that this portion of the analysis does not address the accuracy of the proposed normalization approach, which we have already established, but it examines approaches for reducing the initial residual error.

A simple approach to address the residual error is to up-scale the radar data. The potential gain of this is summarized

in Fig. 9. Error statics presented in this figure were computed in two different ways: initially, we randomly identified ten relatively homogenous fields (based on ground conditions), where the crop type was limited to agricultural vegetation. All ten fields were located along the 40° incidence angle line of flight line 24204, which is in the center of the Kenaston domain. Each field was randomly selected to have a size of 26×41 UAVSAR pixels. The original and normalized RMSE and bias were first computed per field using all available points. The ten error estimates were then averaged, and the resulting values are shown with square symbols in Fig. 9. The same error statistics were computed again, but this time using field average backscatter values (circular symbols). As expected, the up-scaling had no impact in terms of bias; the $\text{Bias}_{\text{HIST}}$ values are consistent with the results from our previous analysis. However, the up-scaling noticeably improved the total $\text{RMSE}_{\text{HIST}}$. The final overall error was reduced to as low as 0.5 dB (see Fig. 9, plot 1.a), which clearly demonstrates the benefit of aggregating the radar data to a coarser grid spacing. Note this reduction is consistent with the square root of the number of looks predicted by (9).

From these analyses, we can summarize that the proposed normalization approach adequately adjusts for incidence angle throughout the full swath width and successfully reduces the systematic differences caused by the incidence angle variability to 0 dB, while the residual component of the total RMSE can be minimized by up-scaling the data, i.e., taking additional looks by spatial averaging.

V. CONCLUSION

Field campaigns with aircraft-based SMAP simulators are needed to develop and validate soil moisture and freeze-thaw retrieval algorithms. A valuable resource, particularly for the radar-based algorithms, is the UAVSAR. UAVSAR collects data over a wide range of incidence angles, and in order to fully exploit it for SMAP, the data must be normalized to a constant incidence angle of 40° . A robust correction technique needs to be able to adequately adjust from any incidence angle to any reference angle irrespective of instrument characteristics, vegetation cover and ground conditions, i.e., to be system and site independent, and to be easily transferable. Here, we normalized the multi incidence UAVSAR data to the SMAP angle of incidence of 40° using a histogram matching procedure, where the adjustment was based on the lowest two central moments, the mean and variance. The HISTogram technique proposed and tested here is similar in terms of theoretical

background to the work presented in [23] and [24], where the normalization was carried out using cumulative frequencies. Thus, as expected, the accuracy performance of both techniques was very similar. The main difference between the two techniques is that in the CDF-based case in the process of “scaling,” the normalized lines will be “forced” to have the same dynamic range as the range of the reference line. The impact of preserving the natural signal variability as opposed to limiting it to the variability observed in the reference line can only be evaluated if a full soil moisture retrieval and comparisons are carried out, which was out of the scope of this paper.

The data set used for analysis was collected during CanEx-SM10. The UAVSAR flight lines were designed so that the whole study area would be covered at a narrow $\sim 10^\circ$ incidence angle range centered at 40° , which consequently also provided large overlap with other flight lines. This meant that only a small portion of the full UAVSAR 22 km swath width would be used when generating the domain composite map. However, due to this large overlap, we were able 1) to evaluate our normalized results against actual data observed at 40° , 2) to explore “alternative” mosaic scenarios generated using fewer number of flight lines, and 3) to test the sensitivity of the HIST technique to different reference angles.

Initial statistics (RMSE_{ORG} and Bias_{ORG}) computed over Kenaston (agriculture-dominated) and BERMS (forest-dominated) between the non-corrected backscatter and the data observed at 40° indicated that the statistics were a function of the angular distance from the reference line. The bias was positive when $\theta_i > \theta_{\text{REF}}$ and negative when $\theta_i < \theta_{\text{REF}}$, and HV-polarized data and forest covered areas showed the least dependence on incidence angle variability. Along with 40° , we tested the robustness of the normalization technique by switching the reference angle to alternative angles (30° and 50°). It should be noted that an ideal model should be able to reduce the bias to 0 dB and lower the total RMSE to the $\text{RMSE}_{\text{res,ORG}}$, which was $2.94 (\pm 0.3)$ dB and $2.3 (\pm 0.1)$ dB for Kenaston and BERMS, respectively. Over Kenaston, for all three reference angle scenarios, the bias component of the total RSME was minimized to 0 dB, and the total $\text{RMSE}_{\text{HIST}}$ was improved and consistently lowered to the level of the initial residual error ($\text{RMSE}_{\text{HIST}} = \sim 3$ dB) for all polarizations, where the variability among the three polarization was < 0.3 dB. Similar $\text{RMSE}_{\text{HIST}}$ results were achieved over BERMS ($\text{RMSE}_{\text{HIST}} = 2.6$ dB (HH), 2.61 dB (HV), 2.52 dB (VV) at $\theta_{\text{REF}} = 40^\circ$). The total $\text{RMSE}_{\text{HIST}}$, which after correction is dominated by the random errors and the residual amplitude bias caused by the additional impact of moisture and roughness variability within the 1° incidence angle lines can be reduced farther by up-scaling the radar data to coarser grid spacing. Overall, it was observed that the accuracy of the histogram-based approach was not sensitive to polarization, choice of reference angle, initial incidence angle, ground conditions in terms of vegetation cover and soil moisture.

As mentioned, the angular uncorrected data are negatively biased relative to the reference data if the original incidence angle is lower than the reference angle and the opposite if the reference angle is lower than θ_i . Bias increases as an expo-

nential function with angular distance. Some of the available normalization approaches do not minimize the bias evenly throughout the swath and as a result the residual bias causes the near- and far-range of the individual swaths/sub-sets to remain obvious (i.e., Fig. 4, plot (c), COS-based normalization). Even though it appears that for $\Delta\theta < 10^\circ$ the angular adjustment is minimal (Fig. 6), under the CanEx-SM10 conditions the overall variability along this narrow strip was significant enough to cause stripping. This problem was solved when the composite maps were generated using the HIST-normalized flight lines. Additionally, the fact that the $\text{RMSE}_{\text{HIST}}$ was consistent (~ 3 dB) irrespective of the original θ_i (Fig. 6) and reference angle supported the hypothesis that a much larger portion of the swath can be used instead of the initially intended 10° . This was tested by generating “alternative” composite maps, which were created using fewer flight lines. The two composite scenarios, one generated using four and other using three out of the seven available flight lines, were evaluated against the available 40° reference data as well as compared against the original seven-lines composite. Error analyses and visual inspection confirmed the similarity. However, the RMSE difference was above 3 dB over approximately 10 (15)% of the area [4 (3)-lines composite]. This can be attributed to differences in reference histograms, urban, or water areas, and impact of row orientation in the agricultural fields. These percentages can be reduced by further subdividing the agricultural class based on row orientation (if present), which is expected to significantly improved the histograms, and/or by implementing a more comprehensive mosaic approach.

The HIST technique presented here is easy to apply and is able to adequately correct the full aircraft swath width. As demonstrated, its accuracy and performance are not dependent on sensor characteristics and/or ground conditions, which indicates its robustness and potential for transferability to other systems and regions. Thus, along with providing data that meets SMAP configuration requirements that can be used for algorithm development and testing, the technique may be beneficial in terms of reducing the number of flight lines, which would eventually result in more cost-effective field campaigns.

ACKNOWLEDGMENT

The U.S. Department of Agriculture (USDA) prohibits discrimination in all its programs and activities on the basis of race, color, national origin, age, disability, applicability, sex, marital status, familial status, parental status, religion, sexual orientation, genetic information, political beliefs, and reprisal or because all or part of an individual's income is derived from any public assistance program (not all prohibited bases apply to all programs). Persons with disabilities who require alternative means for communication of program information (Braille, large print, audiotope, etc.) should contact USDA's TARGET Center at (202) 720-2600 (voice and TDD). To file a complaint of discrimination, write to USDA, Director, Office of Civil Rights, 1400 Independence Avenue, S.W., Washington, D.C. 20250-9410 USA, or call (800) 795-3272 (voice) or (202) 720-6382 (TDD). USDA is an equal opportunity provider and employer.

REFERENCES

- [1] D. Entekhabi, E. Njoku, P. O'Neill, K. H. Kellogg, W. Crow, W. N. Edelstein, J. K. Entin, S. D. Goodman, T. Jackson, J. Johnson, J. S. Kimball, J. R. Piepmeier, R. Koster, N. Martin, K. C. McDonald, M. Moghaddam, S. Moran, R. Reichle, M. Spencer, S. W. Thurman, L. Tsang, and J. Van Zyl, "The Soil Moisture Active Passive (SMAP) mission," *Proc. IEEE*, vol. 98, no. 5, pp. 704–716, May 2010.
- [2] U. Narayan, V. Lakshmi, and E. G. Njoku, "Retrieval of soil moisture from passive and active L/S band sensor (PALS) observations during the soil moisture experiment in 2002 (SMEX02)," *Remote Sens. Environ.*, vol. 92, no. 4, pp. 483–496, Sep. 2004.
- [3] E. G. Njoku, W. J. Wilson, S. H. Yueh, S. J. Dinardo, F. K. Li, T. J. Jackson, V. Lakshmi, and Bolten J., "Observations of soil moisture using a passive and active low-frequency microwave airborne sensor during SGP99," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 12, pp. 2659–2673, Dec. 2002.
- [4] R. Bindlish, T. Jackson, R. Sun, M. Cosh, S. Yueh, and S. Dinardo, "Combined passive and active microwave observations of soil moisture during CLASIC," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 644–648, Oct. 2009.
- [5] N. Chamberlain, M. Zawadzki, G. Sadowy, E. Oakes, K. Brown, and R. Hodges, "The UAVSAR phased array aperture," in *Proc. IEEE Aerosp. Conf.*, Big Sky, MT, Mar. 4–11, 2006.
- [6] S. Hensley, K. Wheeler, G. Sadowy, C. Jones, S. Shaffer, H. Zebker, B. Heavey, E. Chuang, R. Chao, K. Vines, K. Nishimoto, J. Prater, B. Carrico, N. Chamberlain, J. Shimada, M. Simard, B. Chapman, R. Muellerschoen, C. Le, T. Michel, G. Hamilton, D. Robison, G. Neumann, R. Meyer, P. Smith, J. Granger, P. Rosen, D. Flower, and R. Smith, "The UAVSAR instrument: Description and first results," in *Proc. IEEE RADAR*, 2008, pp. 1–6.
- [7] M. C. Dobson, F. T. Ulaby, T. LeToan, A. Beaudoin, E. Kasischke, and N. Christensen, "Dependence of radar backscatter on coniferous forest biomass," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 2, pp. 412–415, Mar. 1992.
- [8] J. Töyrä, A. Pietroniro, and L. W. Martz, "Multisensor hydrologic assessment of a freshwater wetland," *Remote Sens. Environ.*, vol. 75, no. 2, pp. 162–173, Feb. 2001.
- [9] N. Baghdadi, N. Holah, and M. Zribi, "Soil moisture estimation using multi-incidence and multi-polarization ASAR data," *Int. J. Remote Sens.*, vol. 27, no. 10, pp. 1907–1920, May 2006.
- [10] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive. Volume II—Radar Remote Sensing and Surface Scattering and Emission Theory*. Norwood, MA: Artech House, 1982.
- [11] Y. Gauthier, M. Bernier, and J.-P. Fortin, "Aspect and incidence angle sensitivity in ERS-1 SAR data," *Int. J. Remote Sens.*, vol. 19, no. 10, pp. 2001–2006, 1998.
- [12] W. Wagner, J. Noll, M. Borgeaud, and H. Rott, "Monitoring soil moisture over the Canadian Prairies with the ERS scatterometer," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 1, pp. 206–216, Jan. 1999.
- [13] A. Loew, R. Ludwig, and W. Mauser, "Derivation of surface soil moisture from ENVISAT ASAR wide swath and image mode data in agricultural areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 4, pp. 889–899, Apr. 2006.
- [14] R. Magagi, A. Berg, K. Goita, S. Belair, T. Jackson, B. Toth, A. Walker, H. McNairn, P. O'Neill, M. Moghaddam, I. Gherboudj, A. Colliander, M. Cosh, M. Burgin, J. B. Fisher, S. Kim, I. Mladenova, N. Djamaï, L. B. Rousseau, J. Belanger, J. Shang, and A. Merzouki, "Canadian experiment for soil moisture in 2010 (CanEX-SM10): Overview and preliminary results," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 347–363, Jan. 2013.
- [15] R. B. Clapp, "A theoretical and experimental study of radar ground return," MIT Radiation Lab., Cambridge, MA, Rep. 6024, 1946.
- [16] J. P. Ardila, V. Tolpekin, and W. Bijker, "Angular backscatter variation in L-band ALOS ScanSAR images of tropical forest areas," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 821–825, Oct. 2010.
- [17] Y. Oh, K. Sarabandi, and F. T. Ulaby, "An empirical model and an inversion technique for radar scattering from bare soil surfaces," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 2, pp. 370–381, Mar. 1992.
- [18] P. C. Dubois, J. Van Zyl, and T. Engman, "Measuring soil moisture with imaging radars," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 4, pp. 915–926, Jul. 1995.
- [19] J. Shi, J. Wang, A. Y. Hsu, P. E. O'Neill, and E. T. Engman, "Estimation of bare surface soil moisture and surface roughness parameter using L-band SAR image data," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 5, pp. 1254–1266, Sep. 1997.
- [20] K. J. Ranson and G. Sun, "Northern forest classification using temporal multifrequency and multipolarimetric SAR images," *Remote Sens. Environ.*, vol. 47, no. 2, pp. 142–153, Feb. 1994.
- [21] P. J. Hardin and M. W. Jackson, "Investigating SeaWinds terrestrial backscatter: Equatorial savannas of South America," *Photogramm. Eng. Remote Sens.*, vol. 69, no. 11, pp. 1243–1254, 2003.
- [22] F. Baup, E. Mougin, P. Hiernaux, A. Lopes, P. De Rosnay, and I. Chenerie, "Radar signatures of sahelian surfaces in mali using ENVISAT-ASAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 7, pp. 2354–2363, Jul. 2007.
- [23] C. H. Menges, G. J. Hill, W. Ahman, and J. J. van Zyl, "Incidence angle correction of AIRSAR data to facilitate land-cover classification," *Photogramm. Eng. Remote Sens.*, vol. 67, no. 4, pp. 479–490, Apr. 2001.
- [24] C. H. Menges, J. J. Van Zyl, G. J. E. Hill, and W. Ahmad, "A procedure for the correction of the effect of variation in incidence angle on AIRSAR data," *Int. J. Remote Sens.*, vol. 22, no. 5, pp. 829–841, 2001.
- [25] J. J. van Zyl, B. D. Chapman, P. Dubois, and J. Shi, "The effect of topography on SAR calibration," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 5, pp. 1036–1043, Sep. 1993.
- [26] I. Mladenova and V. Lakshmi, "Terrain: Slope influence on quikSCAT backscatter," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2722–2732, Aug. 2009.
- [27] D. Entekhabi, R. H. Reichle, R. D. Koster, and W. T. Crow, "Performance metrics for soil moisture retrievals and application requirements," *J. Hydrometeorol.*, vol. 11, no. 3, pp. 832–840, Jun. 2010.
- [28] A. Toure, K. P. B. Thomson, G. Edwards, R. J. Brown, and B. G. Brisco, "Adaptation of the MIMICS backscattering model to the agricultural context-wheat and Canola at L and C bands," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 1, pp. 47–61, Jan. 1994.
- [29] C. J. Willmott, "Some comments on the evaluation of model performance," *Bull. Amer. Meteorol. Soc.*, vol. 63, no. 11, pp. 1309–1369, 1982.
- [30] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive. Vol. III. From Theory to Applications*. Norwood, MA: Artech House, 1986.
- [31] Y. Wang, J. L. Day, F. W. Davis, and J. M. Melack, "Modeling L-band radar backscatter of Alaskan boreal forest," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 6, pp. 1146–1154, Nov. 1993.
- [32] F. K. Hare, "Climate and zonal divisions of the boreal forest formation in eastern Canada," *Geograph. Rev.*, vol. 40, no. 4, pp. 615–635, Oct. 1950.
- [33] K. J. Ranson, S. Saatchi, and G. Sun, "Boreal forest ecosystem characterization with SIR-C/XSAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 4, pp. 867–876, Jul. 1995.
- [34] S. J. Steele, S. T. Gower, J. G. Vogel, and J. M. Norman, "Root mass, net primary production and turnover in aspen, jack pine and black spruce forests in Saskatchewan and Manitoba, Canada," *Tree Physiol.*, vol. 17, no. 8/9, pp. 577–587, 1997.
- [35] L. L. Bourgeau-Chavez, E. S. Kasischke, and M. D. Rutherford, "Evaluation of ERS SAR data for prediction of fire danger in a Boreal region," *Int. J. Wildland Fire*, vol. 9, no. 3, pp. 183–194, Jan. 1999.
- [36] M. Moghaddam, S. Saatchi, and R. H. Cuenca, "Estimating subcanopy soil moisture with radar," *J. Geophys. Res.*, vol. 105, no. D11, pp. 14 899–14 911, 1998.
- [37] K. Abbott, B. Leblon, G. Staples, M. E. Alexander, and D. MacLean, "Use of RADARSAT-1 images to map forest fuel moisture over boreal forests," in *Proc. IGARSS*, 2002, vol. 1, pp. 134–136.
- [38] M. Whitt and F. Ulaby, "Radar response of periodic vegetation canopies," *Int. J. Remote Sens.*, vol. 15, no. 9, pp. 1813–1848, 1994.



Iliana E. Mladenova received the M.S. degree in hydrology from Free University, Amsterdam, the Netherlands, in 2006, and the Ph.D. degree from the University of South Carolina, Columbia, in 2009. Currently, she is a Research Associate with the U.S. Department of Agriculture, Agricultural Research Service, Hydrology and Remote Sensing Laboratory, Beltsville, MD. Her research interests include development and improvement of passive- and active-based soil moisture retrieval algorithms, and use of microwave remote sensing for improved hydrologic and agricultural prediction.



Thomas J. Jackson (SM'96–F'02) received the Ph.D. degree from the University of Maryland, College Park, in 1976.

He is a Research Hydrologist with the U.S. Department of Agriculture, Agricultural Research Service, Hydrology and Remote Sensing Laboratory, Beltsville, MD. His research involves the application and development of remote sensing technology in hydrology and agriculture, primarily microwave measurement of soil moisture. He is or has been a Member of the science and validation teams of the

Aqua, ADEOS-II, Radarsat, Oceansat-1, Envisat, Advanced Land Observing Satellite, Soil Moisture and Ocean Salinity, Aquarius, Global Change Observation Mission-Water, and Soil Moisture Active Passive remote sensing satellites.

Dr. Jackson is a Fellow of the Society of Photo-Optical Instrumentation Engineers, the American Meteorological Society, and the American Geophysical Union. In 2003, he received the William T. Pecora Award (NASA and Department of Interior) for outstanding contributions toward understanding the earth by means of remote sensing and the AGU Hydrologic Sciences Award for outstanding contributions to the science of hydrology. He received the IEEE Geoscience and Remote Sensing Society Distinguished Achievement Award in 2011.



Scott Hensley (M'05–SM'11) received the B.S. degrees in mathematics and physics from the University of California, Irvine, Irvine and the Ph.D. degree in mathematics from Stony Brook University, Stony Brook, NY, where he specialized in the study of differential geometry.

In 1992, he joined the staff of the Jet Propulsion Laboratory (JPL), Pasadena, where he studies advanced radar techniques for geophysical applications. He has worked on most of the synthetic aperture radar systems developed at JPL over the past two decades including the Magellan and Cassini radars. He was the GeoSAR Chief Scientist, a simultaneous X-band and P-band airborne radar interferometer for mapping above and beneath the canopy that is now commercially operated by Earthdata International. He led the Shuttle Radar Topography Mission Interferometric Processor Development Team for a shuttle-based interferometric radar used to map the earth's topography between $\pm 60^\circ$ latitude. Recently, he began working with the earth-based Goldstone Solar System Radar to generate topographic maps of the lunar surface. He was a Principal Investigator and is currently the Chief Scientist for the NASA UAVSAR program which is an electronically scanned active array L-band fully polarimetric designed for repeat pass radar interferometric applications.



Rajat Bindlish (SM'05) received the B.S. degree in civil engineering from the Indian Institute of Technology, Bombay, India, in 1993, and the M.S. and Ph.D. degrees in civil engineering from The Pennsylvania State University, University Park, in 1996 and 2000, respectively.

Currently, he is with Science System and Applications Inc., working at the U.S. Department of Agriculture Agricultural Research Service, Hydrology and Remote Sensing Laboratory, Beltsville, MD. His research interests involve the application of mi-

crowave remote sensing in hydrology. He is currently working on soil moisture estimation from microwave sensors and their subsequent application in land surface hydrology.